dbSNP: Database of Short Genetic Variations

Catalog of nucleotide changes for human and other model organisms

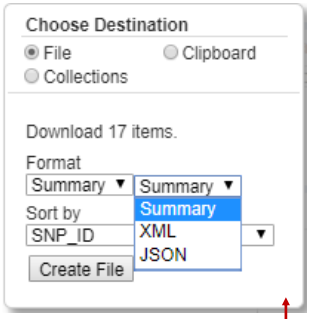
https://www.ncbi.nlm.nih.gov/snp/

National Center for Biotechnology Information • National Library of Medicine • National Institutes of Health • Department of Health and Human Services

Scope and Access

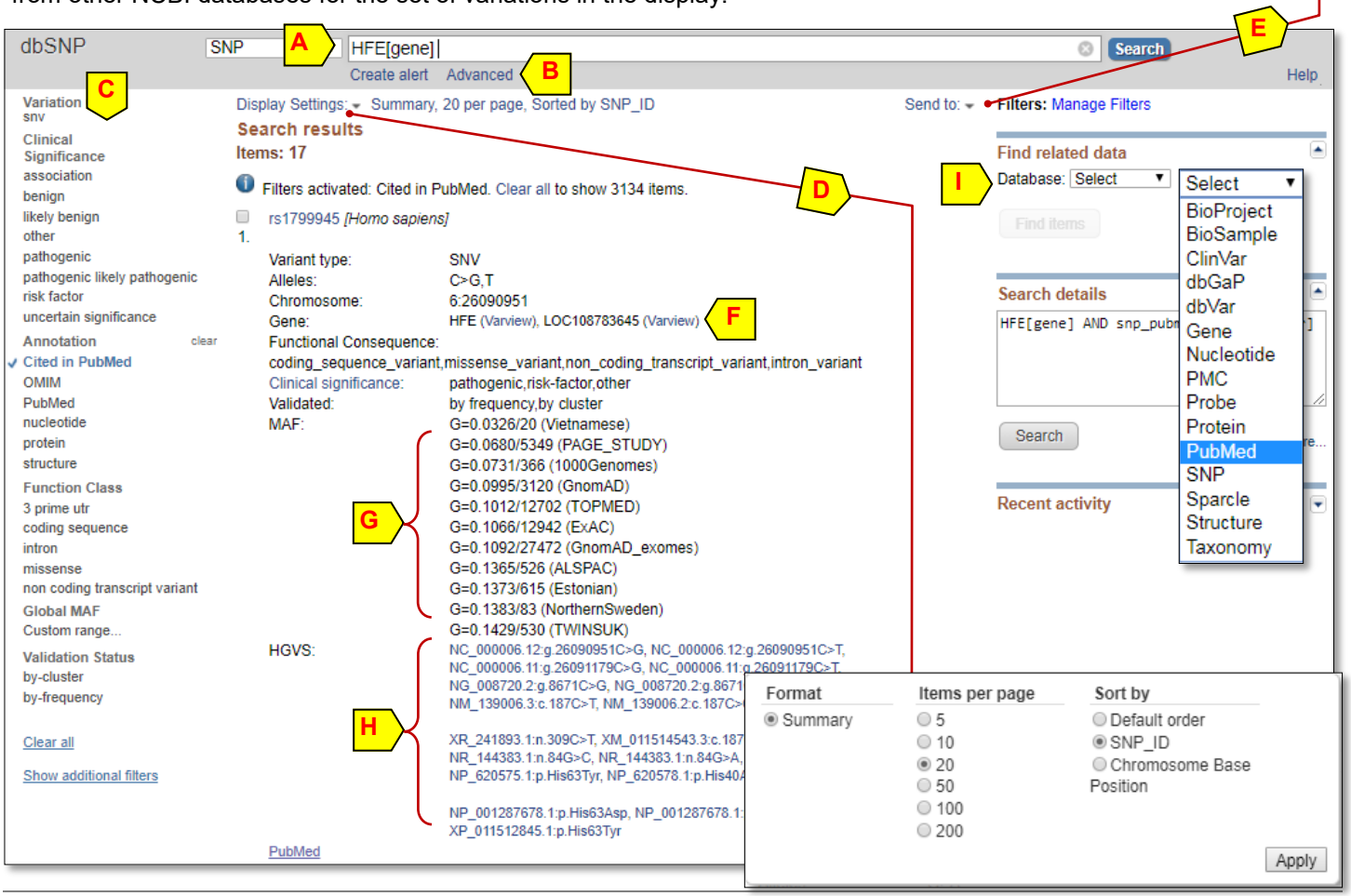
The NCBI Short Genetic Variation database (dbSNP) [1], commonly known as dbSNP, catalogs short variations in nucleotide sequences for human. These variations include single nucleotide variations, as well as insertions, deletions, and short tandem repeats less than 50 nucleotides in length. Short genetic variations may be common, thus representing true polymorphisms, or they may be rare. Some rare human entries have additional information associated with them, including disease associations from ClinVar [2], genotype information and allele origin, as some variations arises in somatic rather than from germline.

Short nucleotide variation data can be accessed via the dbSNP homepage and EUtils API:
www.ncbi.nlm.nih.gov/snp and www.ncbi.nlm.nih.gov/books/NBK25501
VCF files and database bcp files are available for download through FTP and Aspera client at:
ftp.ncbi.nlm.nih.gov/snp/ and www.ncbi.nlm.nih.gov/public/?snp/organisms/
The dbSNP API service, SPDI [3], is available at: api.ncbi.nlm.nih.gov/variation/v0/
dbSNP data can also be accessed interactively through the Variation Viewer:
www.ncbi.nlm.nih.gov/variation/view/



Searching for and Displaying SNP Records

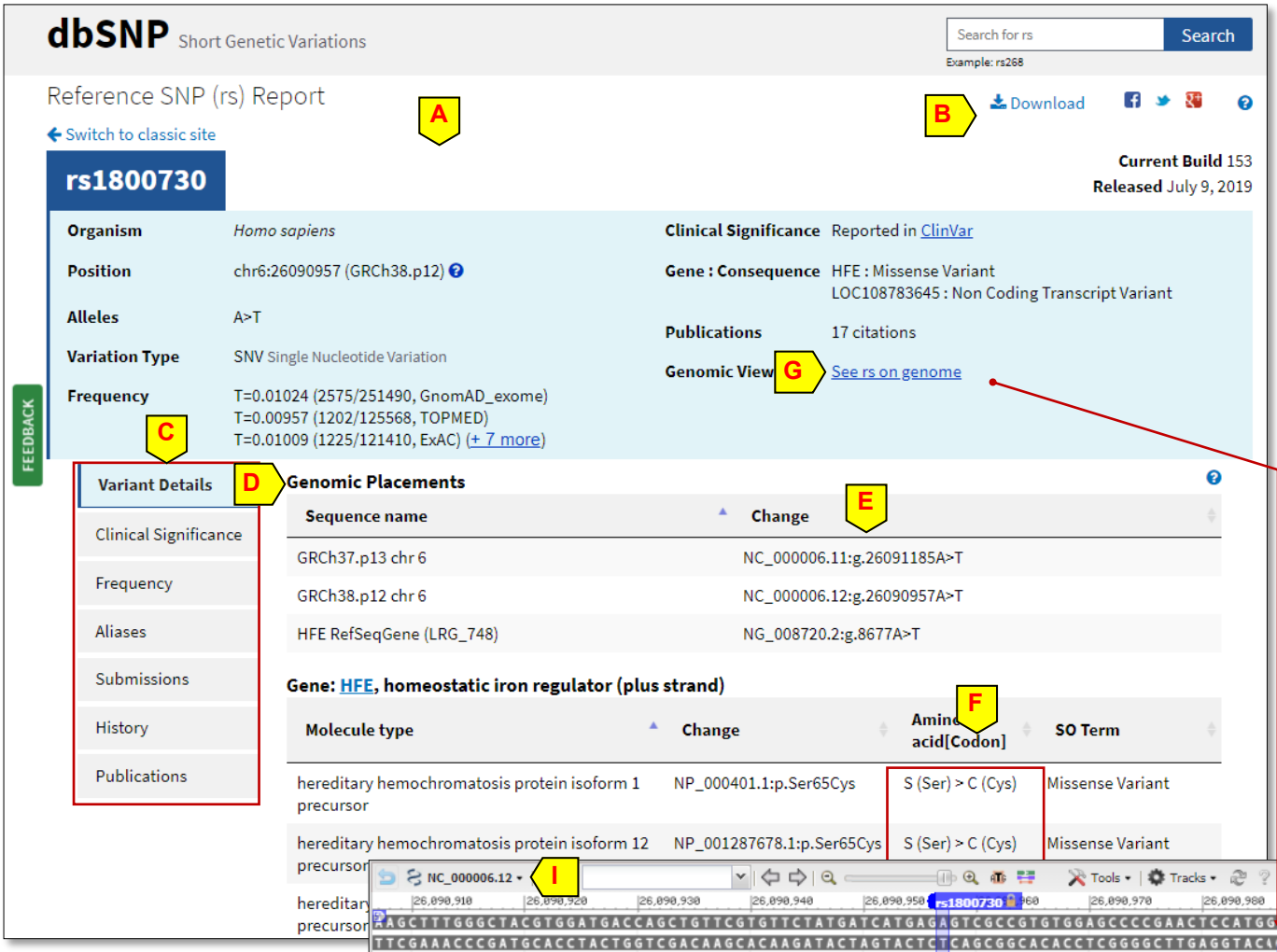
You can search for variations on the dbSNP homepage by typing a query term in the search box and clicking the **Search** button (**A**), or use the **Advanced** (**B**) page to create complex queries to produce more precise results. Searching with hfe[gene] retrieves variations mapped to the human HFE gene, and filter the list to subset using preset filters (**C**). Options in the **Display Settings** popup (**D**) allows you to change the number of records displayed and their sort order. The **Send to** dialog box (**E**) provides options to save retrieved list of SNP to a local file. The **VarView** (**F**) links to a graphical presentation of the variant under the context of the annotated genome in the Variation Viewer. The summary also provides allele frequencies from large population studies (**G**). The **HGVS** variant names (**H**) are hyperlinked to the graphical presentation of the variant on the target molecule presented in Graphical Sequence Viewer. The **Find related data** portlet (**I**) allows you to retrieve related entries from other NCBI databases for the set of variations in the display.



Contact: info@ncbi.nlm.nih.gov

The New Reference SNP Report

The new Reference SNP Report linked from rsIDs, such as rs1800730 (shown below and on p.3) shows details of a dbSNP variation record. The summary section at the top (**A**) provides an overview of the variant. It reports the allele in the forward orientation of the chromosome record. The information in display is also available in JSON format through the Download link at the upper right (**B**). The new report separates details of the variation into various categories (**C**) and displays them next to separate tabs.



The Variant Details tab (**D**), selected by default, displays the genomic placement of the variant on the current and previous genome assemblies (**E**), with the gene and transcript mapping plus the relevant molecular consequences (**F**) shown in the Gene table.

The **See rs# on genome** links to the graphical panel (**G**) of the display. It shows the variant in the context of genome and other neighboring variants. Different tracks groups mapped variants according to specific attributes, such those with literature citation (**H**). Click the link at upper left (**I**) to pop out the Display in a new window.

The New Reference SNP Report (cont.)

Allele: T (allele ID: 15050) **A**

ClinVar Accession	Disease Names	Clinical Significance
RCV000000028.9	Hemochromatosis type 1	Uncertain-Significance
RCV000290779.4	Hereditary hemochromatosis	Uncertain-Significance
RCV000764641.1	Alzheimer's disease,Familial porphyria cutanea tarda,Hemochromatosis type 1,Microvascular complications of diabetes 7,Transferrin serum level quantitative trait locus 2,Variegate porphyria	Uncertain-Significance

Other tabs in the new Reference SNP Report provide category-specific information.

The **Clinical Significance** tab (**A**) lists related clinical assertions for the variant from ClinVar, with IDs linking directly to the records there.

B **C** **D** Download Search: asian

Study	Population	Group	Sample Size	Ref Allele	Alt Allele
1000Genomes	East Asian	Sub	1008	A=1.000	T=0.000
1000Genomes	South Asian	Sub	978	A=1.00	T=0.00
ExAC	Asian	Sub	25164	A=0.9992	T=0.0008
gnomAD - Exomes	Asian	Sub	49008	A=0.9993	T=0.0007
gnomAD - Genomes	East Asian	Sub	1560	A=1.000	T=0.000
The PAGE Study	Asian	Sub	8318	A=1.000	T=0.000
The PAGE Study	SouthAsian	Sub	856	A=1.00	T=0.00

The **Frequency** tab (**B**) lists allele frequency data for major populations from major studies, such as 1000 Genomes, ExAc, and Genome Aggregation Database, with the display filtered for Asian population. Use the "Download" link (**C**) to get the data in a tab-delimited format. This provides a way to evaluate the impact of a variant if there is no information in the **Clinical Significance** and **Publications** sections.

E **F** Search:

Placement	A=	T	Note
GRCh37.p13 chr 6	NC_000006.11:g.260911 85=	NC_000006.11:g.26091185A >T	
GRCh38.p12 chr 6	NC_000006.12:g.260909 57=	NC_000006.12:g.26090957A >T	
HFE RefSeqGene (LRG_748)	NG_008720.2:g.8677=	NG_008720.2:g.8677A>T	
HFE transcript variant 1	NM_000410.3:c.193=	NM_000410.3:c.193A>T	
HFE transcript variant 12	NM_001300749.1:c.193=	NM_001300749.1:c.193A>T	
HFE transcript variant 12	NM_001300749.2:c.193=	NM_001300749.2:c.193A>T	

The **Aliases** tab (**E**) lists HGVS names that all describe the same variant. These names use different reference accessions, but they are equivalent and point to the same genomic variation. Use the table header to sort names or the search box (**F**) to filter the list displayed.

52 SubSNP, 10 Frequency, 3 ClinVar submissions **G** Search:

No	Submitter	Submission ID	Date (Build)
4	1000GENOMES	ss233370756	Jul 14, 2010 (132)
10	1000GENOMES	ss490921024	May 04, 2012 (137)
17	1000GENOMES	ss1319404492	Aug 21, 2014 (142)
53	1000Genomes	NC_000006.11 - 26091185	Oct 12, 2018 (152)
47	ACPOP	ss3733306376	Jul 13, 2019 (153)

H

The **Submission** tab (**G**) lists equivalent submitted entries, from large projects or individual submitters. Only older submissions, before adoption of asserted location, have ssIDs (**H**).

Filter: **I**

Associated ID	History Updated (Build)
rs115372583	Oct 26, 2010 (133)
rs28934888	May 25, 2008 (130)

The **History** tab (**I**) tracks the change of the cluster and lists other rsIDs that have merged with this variant. In this case, variants rs115372583 and rs28934888 were determined to be duplicates of rs1800730, so they were merged into rs1800730 as a single record.

17 citations for rs1800730 **J** Search:

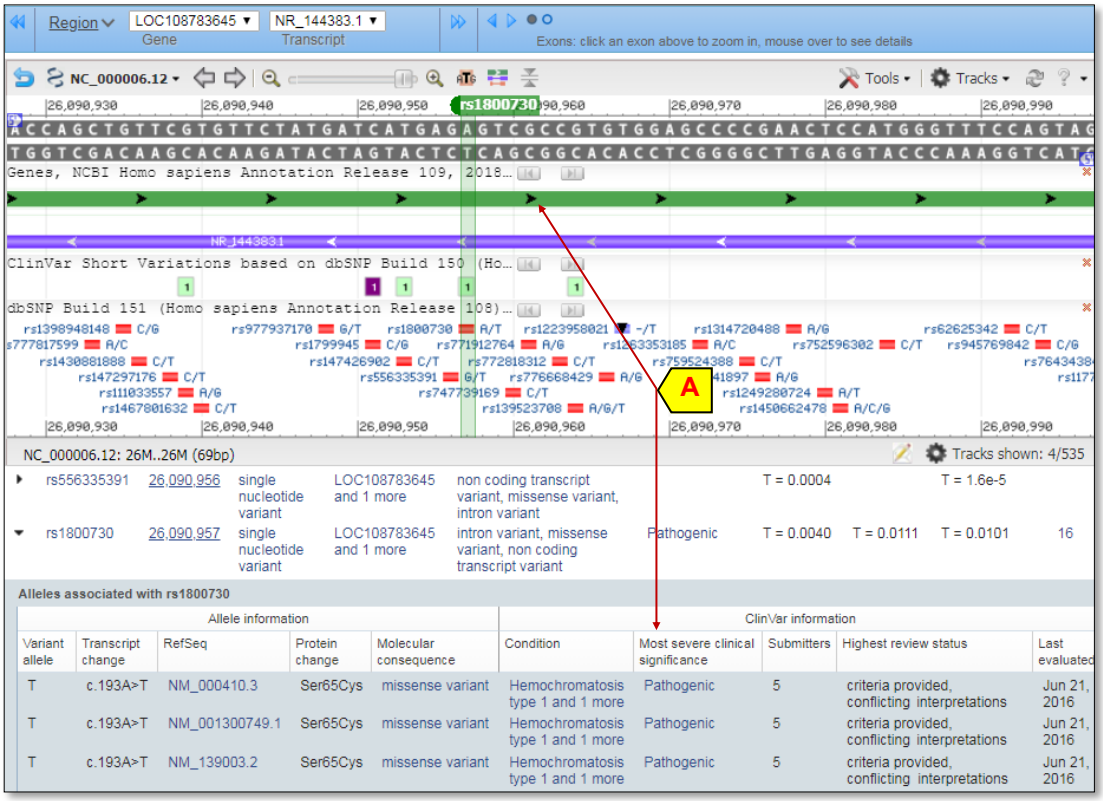
PMID	Title	Author	Year	Journal
30798813	Hemochromatosis: Hereditary hemochromatosis and HFE gene.	Katsarou MS et al.	2019	Vitamins and hormones
26153218	EMQN best practice guidelines for the molecular genetic diagnosis of hereditary hemochromatosis (HH).	Porto G et al.	2016	European journal of human genetics
27173269	Molecular epidemiology of HFE gene polymorphic variants (C282Y, H63D and S65C) in the population of Espírito Santo, Brazil.	Alves LN et al.	2016	Genetics and molecular research
27221532	Population-based analysis of the frequency of HFE gene polymorphisms: Correlation with the susceptibility to develop hereditary hemochromatosis.	Katsarou MS et al.	2016	Molecular medicine reports
27317329	Haplotype analysis of the HFE gene among populations of Northern Eurasia, in patients with metabolic disorders or stomach cancer, and in long-lived people.	Mikhailova SV et al.	2016	BMC genetics

View All in PubMed **K**

dbSNP establishes explicit connections between Reference SNP variants and biomedical literature citations through text-mining. The Reference SNP Report displays these connections under the **Publications** tab (**J**). You can use the "View All in PubMed" button (**K**) to retrieve the list of citations in PubMed and examine their abstracts for more information. Some of these citations may also have free full-text available from PubMed Central (PMC).

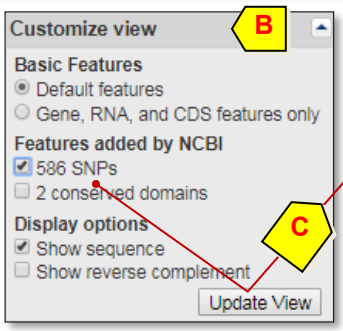
Variation Viewer

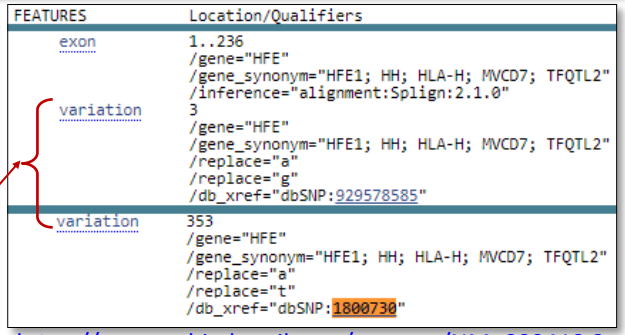
The Variation Viewer provides an interactive display of the variant under the context of annotation of the selected genome assembly. It correlates a variation and its molecular consequences in the data table with its genomic context in the graphical display (**A**). Filters in the left hand column (not shown) are available to selectively display variants of interest. More information on this tool is available online [4, 5]



Other Ways to Access dbSNP Data

The dbSNP database is fully integrated with the Entrez system, enabling the access of variation data through links present in records from other NCBI databases. For example, you can show variations mapped to a RefSeq genomic or mRNA record (with NT_, NG_, NW_ or NM_ accessions) by using the **Customize view** (**B**) menu in the upper right hand corner of the sequence record, simply check the SNPs checkbox and click **Update View** (**C**) to activate the selection.





https://www.ncbi.nlm.nih.gov/nuccore/NM_000410.3

dbSNP also integrates disease-related nucleotide variations that have been reported in literature and cited in rsID format, collected by OMIM, or submitted to ClinVar. The table is the Allelic Variant display for OMIM record 613609, which cites the rsIDs in the dbSNP column (**D**).

613609
HFE GENE; HFE
Allelic Variants (11 Selected Examples) :

http://omim.org/allelicVariant/613609

D

All ClinVar Variants

Number	Phenotype	Mutation	dbSNP	ExAC	ClinVar
.0001	HEMOCHROMATOSIS, TYPE 1 PORPHYRIA CUTANEA TARDA, SUSCEPTIBILITY TO, INCLUDED PORPHYRIA VARIEGATA, SUSCEPTIBILITY TO, INCLUDED HEMOCHROMATOSIS, JUVENILE, DIGENIC, INCLUDED ALZHEIMER DISEASE, SUSCEPTIBILITY TO, INCLUDED TRANSFERRIN SERUM LEVEL QUANTITATIVE TRAIT LOCUS 2, INCLUDED MICROVASCULAR COMPLICATIONS OF DIABETES, SUSCEPTIBILITY TO, 7, INCLUDED	HFE, CYS282TYR	[rs1800562]	-	[RCV000210820...]
.0002	HEMOCHROMATOSIS, TYPE 1 MICROVASCULAR COMPLICATIONS OF DIABETES, SUSCEPTIBILITY TO, 7, INCLUDED	HFE, HIS63ASP	[rs1799945]	[rs1799945]	[RCV000000027...]
.0003	HEMOCHROMATOSIS, TYPE 1	HFE, SER65CYS	[rs1800730]	-	[RCV000290779...]
.0004	HFE INTRONIC POLYMORPHISM	HFE, 5569G-A	[rs1800758]	[rs1800758]	[RCV000000031]
.0005	HFE POLYMORPHISM	HFE, VAL53MET	[rs28934889]	-	[RCV000000032]
.0006	HFE POLYMORPHISM	HFE, VAL59MET	[rs111033557]	-	[RCV000000033]
.0007	HEMOCHROMATOSIS, TYPE 1	HFE, GLN127HIS	[rs28934595]	-	[RCV000000034]
.0008	HEMOCHROMATOSIS, TYPE 1	HFE, ARG330MET	[rs111033558]	-	[RCV000000035]

References

1. The Database of Short Genetic Variation (dbSNP). Kitts A, Phan L, Ward MH, and Holmes JB. In The NCBI Handbook [Internet], 2nd ed. https://www.ncbi.nlm.nih.gov/books/NBK174586/
2. ClinVar: improving access to variant interpretations and supporting evidence. Landrum MJ, et al. Nucleic Acids Res. 2018 Jan 4;46(D1):D1062-D1067. https://www.ncbi.nlm.nih.gov/pubmed/29165669
3. New Web Services for Comparing and Grouping Sequence Variants. https://go.usa.gov/xUeKT.
4. Variation Viewer factsheet. https://ftp.ncbi.nih.gov/pub/factsheets/Factsheet_Variation_Viewer.pdf
5. Variation Viewer Online video tutorial. https://www.youtube.com/watch?v=rnWZ9MFBwUM